Development of Daily Trading Strategies Based on A Quantitative Trading Decision Model

Guangde Shi[1], Jingkai Gao[2], Ruibin Li[3] and Jun Shi[4, *]

[1]Software College, Hefei University of Technology, Hefei, Anhui, 230031

[2]School of Resources and Environmental Engineering, Hefei University of Technology, Hefei, Anhui, 230041

[3]Department of Computer Engineering, Aiyuan Institute of Technology, Taiyuan, Shanxi, 030008

[4]School of Software, Hefei University of Technology, Hefei, 230601

*Corresponding author: juns@hfut.edu.cn

Abstract. Quantitative trading decision models have a key influence on financial investment. Firstly, this study established an LSTM model by using long-term and short-term memory networks and predicted the future prices of gold and bitcoin investment products. Then, according to the time range of gold and bitcoin assets, three types of transactions were determined: cross, non-cross, and inclusion relationship, and the daily trading strategies were determined by the greedy model established by a greedy algorithm. Then, the Sharpe Ratio of the nonparametric method was used to measure the risk of the developed decision model and evaluate the accuracy of the model. Finally, starting from the stock market fluctuation and macro-mobilization, the sensitivity of the decision model under different transaction costs was tested by increasing or decreasing the percentage of transaction costs (0.5%, 1%, 1.5%, and 2%, respectively). Research informs investors on how to invest for the best returns.

Keywords: Trading strategies; Long-term and short-term memory (LSTM) networks; Greedy algorithm; Trading strategies; Quantitative trading decision model.

1. Introduction

People are always looking for suitable investment products and ways to increase non-wage income in real life [1]. Traders often trade unstable assets to maximize revenue [2]. Therefore, starting from current prices, predicting future prices has a very important impact on reducing the volatility of investments. Quantitative transactions are increasingly widely used in the investment field with unique advantages by establishing appropriate mathematical models and formulas to study and analyze the future returns and risks of financial products [3]. In the market investment, extracting and analyzing the key factors that affect the transaction decision is helpful for the investors to obtain the optimal return.

This study aims to develop a quantitative trading decision model to formulate daily trading strategies and simulate actual trading. In addition, the established transaction decision model is tested and verified to prove that the model provides the best transaction strategy, thus providing theoretical support and help for investors to make a more reasonable investment.

2. Style palette

2.1 A quantitative trading decision model

2.1.1. A future price prediction model based on long- and short-term memory network

2.1.1.1. Data preparation and processing

Firstly, the B CHAIN-MKPRU and LBMA-GOLD datasets were used in this study, in which the gold dataset contained 1261 trading days, and the bitcoin dataset contained 1821 trading days. The first 70% of the data in the two data sets were applied as training sets and the last 30% as test sets. In addition, time data needs to be converted into digital data. In the study, the time interval between two adjacent transactions of users was expressed by converting specific time into time, and the conversion of data types was completed.

Secondly, the typical predicted data values were set by classifying the two data sets. The data of every six trading days were organized into a group, and the last group of data of each group was taken as the standard set of predicted data, thus avoiding data flattening. The data set was normalized using the zero-mean normalization method, and the formula is as follows:

$$x^* = \frac{x - \bar{x}}{\sigma} \tag{1}$$

Where x and $\bar{x}$ are a value and the mean value of an original data, respectively, and σ is the standard deviation of an original data.

The formula calculates the standard deviation:

$$\sigma = \frac{1}{\sqrt{N}\sum_{i=1}^{N}\left(x_i - \mu\right)^2} \tag{2}$$

Finally, the loss function used to train the long and short-term memory network (LSTM) was specified as cross-entropy loss [4]. In the LSTM, backpropagation trained the parameters together to reduce their errors.

2.1.1.2. Introduction of the LSTM network

A recurrent neural network (RNN) is a neural network based on a feedforward neural network (FNN), which is different from ordinary networks. The RNN has short-term memory, which can better process data related to the sequence. Therefore, RNN has more advantages in processing sequence-related data. The LSTM solves the problems of gradient disappearance and explosion of RNN and overcomes the problem that RNN cannot handle the high-order dependence of data [5]. The neurons in LSTM contain three gates: input gate, forgetting gate, and output gate. The gate can be regarded as a fully connected layer with a sigmoid function as the activation function [6].

$$g(x) = \sigma(Wx + b) \tag{3}$$

$$\sigma(x) = 1/(1 + \exp(-x)) \tag{4}$$

Gating is achieved through the dot product operation of the sigmoid function and data flow.

2.1.1.3. Parameter setting of the LSTM network

Under the condition that the data set was X as training data and Y as verification data, the parameters of the LSTM network were set as follows: the length of the window (the number of transaction intervals needed to predict the next transaction interval) was 5. The number of input neurons was N1. The number of neurons in the GRU unit (the first layer neural network) was 20. Then, the number of GRU units was 7. Furthermore, the activation functions of the forgetting gate and output gate used the default settings of Python's Keras library, and the learning rate was 0.0006. The rejection rate of each layer of network nodes was 0.2, and the error calculation method used the mean square error; The iterations (epoch and batch size) of determining weight parameters by an RMSProp algorithm [7-8]. In addition, generally speaking, the more layers of LSTM (generally no more than 3 layers, the more training is challenging to converge), the stronger the learning ability of high-level time representation, so an ordinary neural network layer was added to reduce the dimension of output results.

If it needs to train multiple sequences in the same model, it can input these sequences into independent LSTM modules and then combine the output results and input them into the common layer. In addition, the data set was divided into training set and verification set at the ratio of 7:3, which prevents the model from falling into over-fitting training. Then, the X column of data was imported into the model as a parameter to get the predicted value. The performance of the model was evaluated by comparing the error between the predicted value and the actual value.

2.1.1.4. Development of an LSTM model

The training steps of an LSTM model include the forward calculation of data information and reverse transmission of errors, and they are as follows:

Step1: It is necessary to design the network structure to establish an LSTM model and select the required loss function.

Step2: After the model is established, the model's parameters need to be initialized. In addition, the estimated value of the model is solved by forwarding calculation, and the loss function value of the model is calculated according to the estimated value and the actual value.

Step3: The gradient information of the parameters is calculated according to the derivative of the loss function. The model's parameters need to be updated according to the model's learning rate and gradient information, and the new parameters are reintroduced into the model for data forward calculation information.

Step4: Through the second and third steps of the recurrent calculation, the model's parameters are finally determined, and a suitable LSTM model will be established, which can be applied to actual classification or prediction.

2.1.1.5. solution of the LSTM model

After the LSTM model was developed, it was found that the speed and volatility of the price change trend of gold and bitcoin had an important influence on predicting their future price changes. Figure 1 and Figure 2 show the changes of the predicted (blue polyline) and actual (solid red line) values of bitcoin and gold prices, respectively.

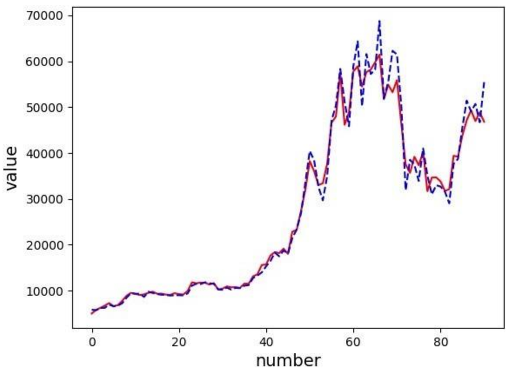

Fig. 1 Changes of the predicted and actual values of bitcoin prices

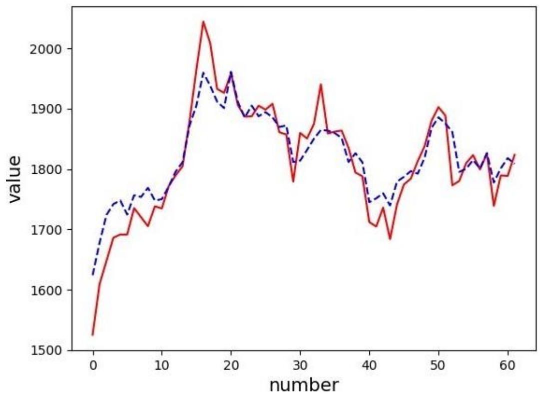

Fig. 2 Changes of the predicted and actual values of gold prices

The results of Figure 1 and Figure 2 show that the LSTM model is not sensitive enough, and there is a lag when the data changes greatly. However, on the whole, the price change predicted by this model is in good agreement with the actual price change, which is of reference significance.

2.1.2. A greedy model

2.1.2.1. Determination of gold and bitcoin transaction types based on the greedy algorithm

In the optimization process in the next few days, the time range of gold and bitcoin assets can be got, which can be divided into three situations: one is the partial intersection, the other is the subset relationship, and there is no intersection.

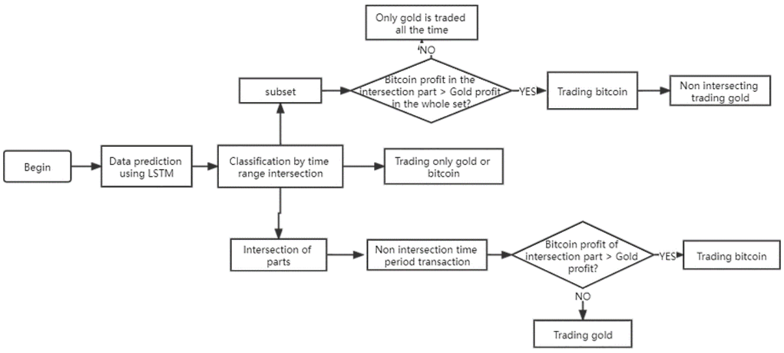

Fig. 3 Process of determining transaction type

Step1: If the return of gold is higher than that of bitcoin in 3-5 days, and because gold is the local optimal solution in 1-5 days and bitcoin has no return in 1-3 days, gold needs to be purchased. Because bitcoin has an optimal solution in 3-7 days, including a subset 5-7 days, it is also profitable to buy bitcoin. Therefore, buying gold in 1-5 days and bitcoin in 5-7 days can achieve the maximum profit.

Step2: If the return of gold is lower than that of bitcoin in 3-5 days since bitcoin is the local optimal solution in 3-7 days and gold has no return in 5-7 days, bitcoin needs to be purchased in 3-7 days. Because gold has an optimal solution in 1-5 days, including subset 1-3 days, buying gold in 1-3 days can also make a profit. To sum up, buying gold in 1-3 days and bitcoin in 3-7 days can get the maximum profit.

Step3: If the return of gold is higher than that of bitcoin in 3-5 days, and bitcoin has an optimal solution in 1-7 days, including subset 1-3 days and 5-7 days, then the profit can be obtained in 1-3 days and 5-7 days. To sum up, buying bitcoin in 1-3 days, buying gold in 3-5 days, and rebuying bitcoin in 5-7 days can get the maximum profit.

Step4: If the return of gold is lower than that of bitcoin in 3-5 days, and there is no other subset of gold in 1-7 days, that is, there is no return, so you do not need to buy gold in 1-7 days, buy bitcoin. To sum up, buying bitcoin in 1-7 days can get the maximum profit.

2.1.2.2. Determination of trading strategies

According to the predicted price data for the next 15 days, the lowest price in the next 15 days can be selected as the starting price and the highest price as the selling price. It is obtained the relative optimal transaction value in the next 15 days. The combination of the optimal relative values for each epoch can form the overall optimal solution.

2.1.3. Risk assessment

2.1.3.1. Nonparametric method of Sharpe Ratio

Sharpe Ratio is an index that can balance profit and risk. It was chosen to measure the risk to return ratio in the study. The calculation method of the Sharpe Ratio index is as follows:

Step1: To establish a heteroscedasticity nonparametric model.

$$y_i = \mu(x_i) + \sigma(x_i)\varepsilon_i \quad i = 1,\ 2,\ \cdots,\ n \tag{5}$$

Step2: To calculate the conditional mean function $\mu(x)$.

To take independent random sample points (x_i, y_i) and use the smooth spline method to calculate the conditional mean function.

$$\mu(x)=\sum_{j=1}^{N}N_j(x)\hat{\theta}_j \tag{6}$$

Where:

$$\hat{\theta}=(N^TN+\lambda\Omega_N)^{-1}N^Ty \tag{7}$$

Step3: To calculate the variance $\sigma^2(x)$.

For $\mu(x)$, the variance is calculated using a residual-based approach.

$$r(x_i)=(y_{i-\mu}(x_i))^2 \tag{8}$$

$$E(y_i(x_i))=E(y_i-\mu(x_i))^2=\sigma^2(x_i) \tag{9}$$

Then, the sample points are taken.

$$(x_i,\ r(x_i)) \tag{10}$$

The variance is obtained by minimizing the objective function using local linear kernel estimation.

$$\begin{aligned}(\hat{\alpha}_i,\hat{\beta}_i)=\arg\min_{\alpha,\beta}\sum_{i=1}^{n}(\hat{r}_i-\alpha-\beta(x_i-x))^2\\ *K_h(x_i-x)\end{aligned} \tag{11}$$

Where:

$$K_h(x_i-x)=\frac{1}{h}K(\frac{x_i-x}{h}),\quad h>0 \tag{12}$$

The appropriate window width parameter is:

$$\hat{\alpha}_i=Er(x_i)=\sigma^2(x_i) \tag{13}$$

Step4: The result of the Sharpe Ratio is obtained by the ratio calculation as follows:

$$f(x_i)=\frac{\mu(x_i)}{\sigma(x_i)} \tag{14}$$

2.1.3.2. Results analysis

The Sharpe Ratio index represents how much investors can get paid for each additional risk point. If the index is greater than 1, the yield is higher than the fluctuation risk; If less than 1, the operational risk is greater than the rate of return. After calculating by intercepting some falling points of the total value, it was found that when the actual price changed greatly, the model would lag. Specifically, when choosing the highest price and lowest price within 15 days, some high-yield predicted data would be quite different from the actual data, leading to mistakes and losses in decision-making.

2.2 Evaluation of the quantitative trading decision model

2.2.1. Accuracy analysis of the model

2.2.1.1. Introduction of the moving average convergence divergence index

The moving average convergence divergence (MACD) is widely used for model comparison and error analysis [9]. The MACD is developed from the double exponential moving average. The convergence and separation between the short-term and long-term moving averages of MACD closing price were used to make buying and selling decisions. The calculation process is as follows:

Step1: To calculate the exponential moving average (EMA).

$$EMA_{today} = \alpha * price_{today} + (1-\alpha) * EMA_{yesterday} \tag{15}$$

Where α is the smoothness index, which is generally taken as 2/(N+1), and N is generally chosen as 12 and 26.

Step2: To calculate the differential value (DIF).

$$\begin{aligned} DIF &= EMA(CLOSE, SHORT) \\ &- EMA(CLOSE, LONG) \end{aligned} \tag{16}$$

Where *SHORT* and *LONG* are generally taken as 12 and 26, respectively.

Step3: To calculate the differential exponential average value (DEA).

$$DEA = EMA(DIF, MID) \tag{17}$$

Where MID is generally taken as 9.

Step4: To calculate the MACD index

$$MACD = 2 * (DIF - DEA) \tag{18}$$

2.2.1.2. Result analysis of the accuracy of the model

In this study, EMA(26) was used to subtract EMA(12) to get DIF, and then the MACD index was calculated by formula (18) when the MID was 9. Figure 4 and Figure 5 show the accuracy analysis results of gold and bitcoin respectively.

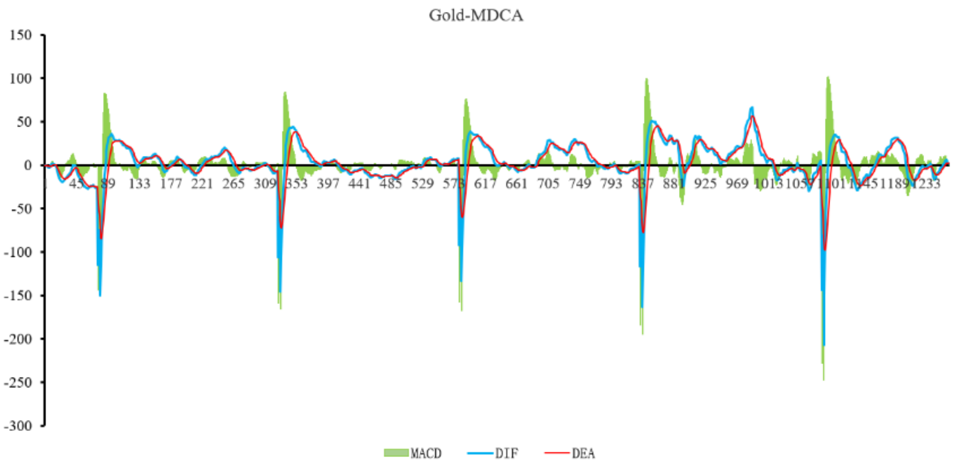

Fig. 4 Accuracy analysis results of gold

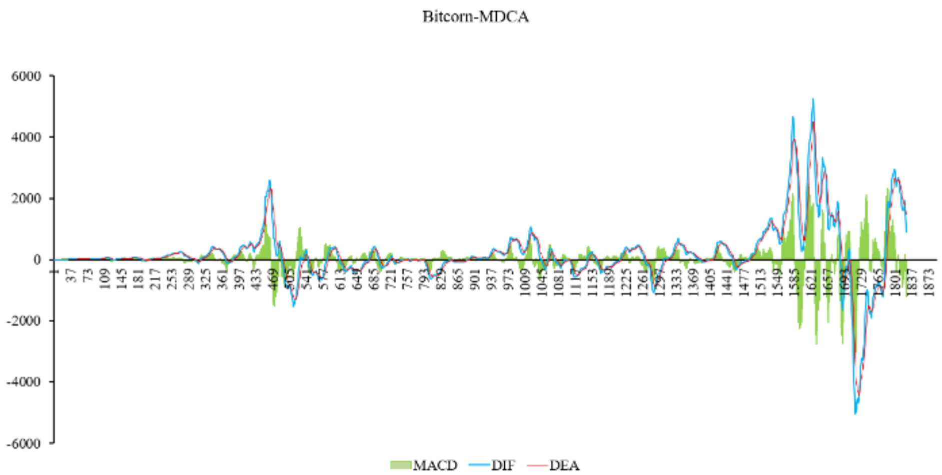

Fig. 5 Accuracy analysis results of bitcoin

2.2.2. Sensitivity analysis of investment strategies to transaction costs

In establishing the above model, the difference in transaction cost may affect the results of prediction and decision-making. Therefore, the sensitivity under different transaction costs (transaction decision results) could be tested by increasing or decreasing the percentage of transaction costs. To rerun the model by increasing or decreasing the transaction cost by 0.5%, 1%, 1.5%, and 2%. The results are shown in Figure 6 below. Figure 6 shows that the model results are better under the same transaction cost.

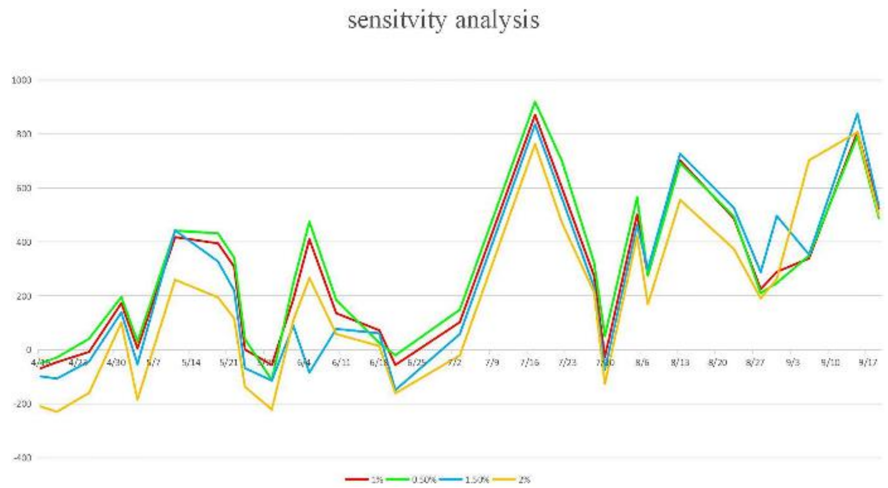

Fig. 6 Variations of the model result with the changes of transaction cost

3. Model assessment

This study formulated the daily trading strategies by establishing a quantitative trading decision model. These advantages are as follows: Firstly, the future price changes were simulated using the long-term and short-term memory networks of an LSTM model. Secondly, according to the predicted price data for the next 15 days, the lowest starting price and the highest selling price were selected to obtain the relative optimal transaction value for the next 15 days. Finally, the accuracy and sensitivity of the model were analyzed by using a Sharpe Ratio index and a MACD index, respectively. Removing high-risk points through Sharpe Ratio could effectively reduce negative returns. At the same time, the performance of the model at different sensitivity levels was analyzed from two aspects of stock market fluctuation or macro mobilization.

However, the established model still has such limitations: the LSTM network is complex and has many parameter settings. In addition, the application of the greedy algorithm often can only get the optimal local solution but cannot get the optimal global solution [10], which is where this research needs to focus on optimization in the future application.

References

[1] X.X. Chen, Research on asset allocation of Chinese investors in global market, Tongji University, Shanghai, 2008.

[2] Y. Li, Research on the application of quantitative trading strategy model, Yunnan University of Finance and Economics, Kunming, 2014.

[3] J.J. Wang, Quantitative trading applications in China's stock market, Nanjing University, Nanjing, 2013.

[4] C. Ravi, Fuzzy Crow Search Algorithm-Based Deep LSTM for Bitcoin Prediction, in: International Journal of Distributed Systems and Technologies (IJDST), IGI Global, Pennsylvania, Hershey, 2020, pp. 53-71.

[5] B.B. Sahoo, R. Jha, A. Singh, D. Kumar, Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting, in: Acta Geophysica, Springer, Berlin, Heidelberg, 2019, pp. 1471–1481.

[6] W. Bao, J. Yue, Y.L, Rao, A deep learning framework for financial time series using stacked autoencoders and long-short term memory, in: PLoS ONE, Public Library of Science, San Francisco, California, 2017, pp. 1–24.

[7] Y.Q. Fang, Z. Lu, J.W. Ge, Forecasting stock prices with a combined RMSE loss LSTM-CNN model, in: Computer Engineering and Applications, Beijing "Computer Engineering and Application" Journal Co., Ltd., Beijing, 2022, pp. 1-10.

[8] S. Mehtab, J. Sen, Stock Price Prediction Using CNN and LSTM-Based Deep Learning Models, in: 2020 International Conference on Decision Aid Sciences and Application (DASA), IEEE Press, Piscataway, NJ, 2020, pp. 447-453.

[9] A.A.A. Aguirre, N.D.D. Méndez, R.A.R. Medina, Artificial intelligence applied to investment in variable income through the MACD (moving average convergence/divergence) indicator, in: Journal of Economics, Finance and Administrative Science, Emerald Publishing, United Kingdom, 2021, pp. 268-281.

[10] T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein, Introduction to Algorithms third edition, The MIT Press, 2009.